# SOUND SCOPE PAD

**Masatoshi Hamanaka**
RIKEN
masatoshi.hamanaka@riken.jp

## ABSTRACT

We developed Sound Scope Pad to provide an active music listening experience that combines AI, virtual reality (VR), and spatial acoustics. Users can emphasize the sounds of different performers by turning their head to the left or right or bringing their hands closer to their ears, allowing them to find and focus on the performer they want to hear. In the Sound Scope Headphones that we previously built, the user's head direction was detected by an accelerometer mounted on the headphones arc. In Sound Scope Pad, we enable detection in the head direction by combining the angle information detected by the acceleration gyro sensor of a tablet and the angle information of the head recognized from the front camera image of the tablet.

## 1. INTRODUCTION

Sound Scope Pad is an application that enables users to focus on the sound of a specific part of a song more clearly while listening to it (Fig. 1). Turning their head to the left or right enables users to follow the sound of specific instruments on the respective sides. Moving their palms close to their ears in a "listening pose" enables users to focus on only specific parts.

We previously proposed Sound Scope Headphones that achieved the aforementioned functions by using a digital compass that detects the orientation of the face and distance sensors that measure the distance between a hand and an ear [1, 2]. Figure 2 shows our exhibition at ACM SIGGRAPH2009 Emerging Technologies. Since face-to-face exhibitions are currently difficult due to social distancing in the wake of COVID-19, we investigated implementing similar functions in an app so that many people can experience it using a tablet.

## 2. RELATED WORKS

Headphones equipped with sensors that detect the direction and position of the head have been available for a while, but since their purpose is to enhance the sense of presence by fixing the virtual sound source position [3–5], they cannot be used to emphasize a specific part. For example, these headphones make it difficult to selectively listen to only a certain instrument located nearby other instruments.

Figure 1. Exhibition of Sound Scope Headphones

The spatial audio system used enables the volume of an instrument to be changed by moving the position of the listener's avatar and each part [6, 7]. However, it can be difficult for beginners to properly adjust the mixing of each part because this requires a complicated operation in which the part where a solo start must be placed in the center localization and the part where it returns to the accompaniment must be placed far away.

## 3. SOUND SCOPE PAD

By installing the Sound Scope Pad application on a tablet, users can experience the following (Fig. 2).

**VR Concert Experience** Performers stand around you and start playing, and you can see three at a time in front of you on your iPad screen. If you point the iPad to the left, the performer will appear on the left, and if you point it to the right, the performer will appear on the right.

**Spatial Acoustic Experience** If you hold the iPad in front of you, you will see the application tracking your face on the display. By turning your head to the left or right, you can follow the sound of the performers on their respective sides. Likewise, if you turn around with the iPad in hand, you will be able to face the performers behind you.

**Active Music Listening Experience** This app provides an active music listening experience by enabling you to focus on certain performers that you want to hear. If you turn to face the performer of your choice, it will sound as though they are directly in front of you. By moving your left or right palm closer or further from your ear, you can adjust the range of the performance that you want to focus on. You can highlight one part or multiple parts of the performance, such as the horn section or the rhythm section. You can check the direction of the performer and the direction you are facing on the position display.
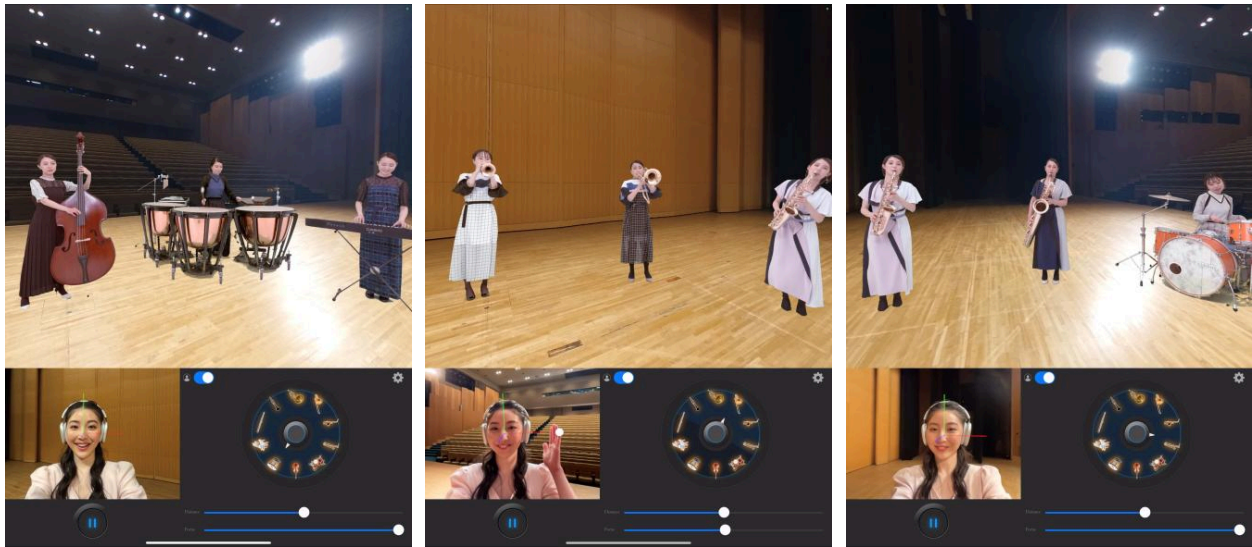
Figure 2. Screen snapshot of Sound Scope Pad.

## 4. IMPLEMENTATION

Let $\theta_n$ be the angle created from each of the *n* parts from the listener's avatar. By changing the amplification rate $h_n^\theta$ ($0 \le h_n^\theta \le 1$) in accordance with the orientation of the head, the sound of the part at an angle close to the front can be more emphasized. In the following, the angle between the head direction and the position where the part is placed is $\theta$ ($-\pi \le \theta < -\pi$), and the distance between the hand and ear is $\delta$ ($0 \le \delta \le 1$).

$$h_n{}^\theta = \begin{cases} 0 & \widetilde{h_n{}^\theta} < 0 \\ \widetilde{h_n{}^\theta} & 0 \ge \widetilde{h_n{}^\theta} \end{cases} \qquad (1),$$

where

$$h_n{}^\theta = \begin{cases} 0 & \widetilde{h_n{}^\theta} < 0 \\ \widetilde{h_n{}^\theta} & 0 \ge \widetilde{h_n{}^\theta} \end{cases} \qquad (2).$$

$\alpha$ ($0 \le \alpha \le 1$) is an adjustable parameter that sets the change in amplification rate when $\delta < 1$. When $\alpha=0$, the amplification rate of each part does not change even if the hand is closer to the ear, but when $\alpha > 0$, the amplification rate decreases as the hand approaches the ear. At this time, since the decrease in the amplification rate is larger in the direction in which the listener is not facing, the sound in front can be heard relatively loudly.

## 5. CONCLUSION

We have described Sound Scope Pad, an application that enables users to emphasize the part they want to listen to in a song consisting of multiple parts by using head direction and hand gestures. The Sound Scope Pad app and an introductory video explaining how to use it can be downloaded at

`https://gttm.jp/hamanaka/en/soundscopepad/`

## 6. REFERENCES

[1] Masatoshi Hamanaka and Sunghee Lee, "Sound Scope Headphones," ACM SIGGRAPH, 2009, Talks TK-201 / Emerging Technologies ET-201.

[2] Masatoshi Hamanaka and Sunghee Lee, "Music Scope Headphones: Natural User Interface for Selection of Music," In Proc. of the 7th Int. Conf. on Music Information Retrieval, Victoria, Canada, 2006, pp. 302–307.

[3] Oliver Warusfel and Gerhard Eckel, "LISTEN: Augmenting Everyday Environments Through Interactive Soundscapes," In Proc. of IEEE Workshop on VR for Public Consumption, IEEE Virtual Reality, 2004, Chicago, USA, pp. 268–275.

[4] Jiann-Rong Wu, Cha-Dong Duh, M. Ouhyoung, and Jei-Tun Wu, "Head Motion and Latency Compensation on Localization of 3D Sound in Virtual Reality," In Proc. of the ACM Symposium on Virtual Reality Software and Technology, ACM Virtual Reality Software and Technology, 1997, Lausanne, Switzerland, pp. 15–20.

[5] Camille Goudeseune and Hank Kaczmarski, "Composing Outdoor Augmented-Reality Sound Environments," In Proc. of the Int. Computer Music Conf., International Computer Music Association, 2001, Havana, Cuba, pp. 83–86.

[6] François Pachet and Oliver Delerue, "A Mixed 2D/3D Interface for Music Spatialization," In Lecture Notes in Computer Science (no. 1434) First Int. Conf. on Virtual Worlds, 1998, Paris, France, pp. 298–307.

[7] François Pachet and Oliver Delerue. "On-The-Fly Multi-Track Mixing," In Proceedings of AES 109th Convention , Audio Engineering Society, 2000, Los Angeles.